Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions: a systematic review

Balzer, Julia ; van der Linden, Marietta L ; Mercer, Thomas H ; van Hedel, Hubertus J A

Abstract: AIM: Recovery and trainability of impaired selective voluntary motor control (SVMC) of the lower extremity in children with upper motor neuron lesions has received little attention. To facilitate an evidence-based debate about this topic, this review evaluates the evidence level of the psychometric properties of SVMC measures. METHOD: MEDLINE, Embase, CINAHL, PsycINFO, Scopus, Cochrane and PEDro databases were systematically searched up to July 2016. Two independent raters scored the methodological quality in accordance to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. The overall level of evidence was scored according to Cochrane criteria. RESULTS: We identified 3590 studies, of which 17 were included. COSMIN scores ranged from 'poor' to 'excellent' for studies investigating measurement properties of the Selective Motor Control test, modified Trost test, Gillette's Selective Motor Control test, Selective Control Assessment of the Lower Extremity (SCALE), kinematic measures, electromyography, and torque steadiness. Studies assessing the SCALE scored highest on COSMIN items. Evidence levels for SCALE's validity and reliability properties were moderate, while for the other SVMC measures these ranged from unknown to moderate. Responsiveness was not assessed. INTERPRETATION: Further psychometric studies of SVMC measures are needed to provide a scientific contribution to the ongoing debate of SVMC trainability.

DOI: https://doi.org/10.1111/dmcn.13417

Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions - A systematic review

Authors

Julia Balzer (MSc PT)
Pediatric Rehabilitation Research Group
Rehabilitation Center Affoltern am Albis
University Children's Hospital Zurich
Mühlebergstrasse 104
CH-8910 Affoltern am Albis
Switzerland
Phone: +41 (0)44 762 52 91
Fax: +41 (0)44 762 51 22
Email: Julia.Balzer@kispi.uzh.ch

Marietta L van der Linden (MSc PhD)
Centre for Health, Activity and Rehabilitation Research, Queen Margaret University – Edinburgh (GB)

Thomas H Mercer (MSc PhD)
Centre for Health, Activity and Rehabilitation Research, Queen Margaret University – Edinburgh (GB)

Hubertus JA van Hedel (PhD PT)
Pediatric Rehabilitation Research Group, Rehabilitation Center Affoltern am Albis, University Children's Hospital Zurich (CH)

Word count **4212**

Tables

Table I: General Characteristic, psychometric properties and clinical utility of SVMC measures
Table II: Summary study details, COSMIN-, quality and evidence-rating

Figures

Figure 1: Selection process systematic review

Appendixes

Appendix 1: Search Construct: Psychometric studies of SVMC measures for the lower extremity in children with UMN lesions
Appendix 2: Quality Criteria for Measurement Properties
Appendix 3: Levels of Evidence for the Overall Quality of the Measurement Properties (Based on the Cochrane Back Review Group 2003
Appendix 4: Specified COSMIN rating rules

Abstract

Aim

Recovery and trainability of impaired selective voluntary motor control (SVMC) of the lower extremity in children with upper motor neuron lesions (UMN) has received little attention. To facilitate an evidence-based debate about this topic, this review evaluates the evidence level of the psychometric properties of SVMC measures.

Method

MEDLINE, EMBASE, CINAHL, PsycINFO and SCOPUS databases were systematically searched up to July 2016. Two independent raters scored the methodological quality in accordance to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. The overall level of evidence was scored according to Cochrane criteria.

Results

We identified 3590 studies of which 17 were included. COSMIN scores ranged from 'poor' to 'excellent' for studies investigating measurement properties of the 'Selective Motor Control (SMC) test', 'modified Trost-test', 'Gillette's SMC test', 'Selective Control Assessment of the lower Extremity' (SCALE), kinematic measures, electromyography, and torque steadiness. Studies assessing the SCALE scored highest on COSMIN items. Evidence levels for SCALE's validity and reliability properties were 'moderate' while for the other SVMC measures these ranged from 'unknown' to 'moderate'. Responsiveness was not assessed.

Interpretation

Further psychometric studies of SVMC measures are needed to provide a scientifically contribution to the ongoing debate of SVMC trainability.

Shorted title: Psychometric properties of SVMC measures

'What this paper adds':

- Very few studies evaluate validity and reliability of SVMC measures
- Responsiveness studies of SVMC measures are lacking
- We can currently recommend the 'SCALE' measure/tool
- The evidence level of objective interval-scaled SVMC measures is limited
- Further rigorous psychometric studies of SVMC measures are needed

Introduction

Understanding and therapeutically guiding motor development and motor control is a complex and challenging topic for professionals and caregivers within the field of neuropediatric rehabilitation.[1-3] Motor control is a primary determinant for moving physiologically or 'normally.[4] Measuring motor control regards the question *'how are motor units selected, activated and deactivated'*.[5] In children with upper motor neuron (UMN) lesions, the muscle control mechanism of selected activation and deactivation is often disturbed and causes non-physiological movements (patterns).[3,6] Clinically, this is known as a loss of selective voluntary motor control (SVMC).[6-8] Selective motor control is defined as the ability 'to isolate the activation of muscles in a selected pattern in response to demands of a voluntary movement or posture'. [8] The term 'voluntary' within SVMC emphasises the deliberate performance of selected muscle activation during functional tasks.[7]

Pathophysiologically, a loss of SVMC is related to impaired descending corticospinal input, which results in disturbed control of spinal networks.[6,8] Complex muscle activation patterns of agonist, synergist, and antagonist are disturbed.[8,9] This allows the appearance of flexor and/or extensor mass movement patterns (i.e. synergies), which often limits the patient's control of movement.[7] For patients with a complete loss of SVMC, strength and functional training is only possible within synergies, potentially accompanied by co- and mirror movements of other muscle groups. This could limit the patient's ability to increase or maintain strength of a specific weak muscle (group).[6-9] Therefore, in the long term, impaired SVMC can result in a vicious circle of limited motor performance during daily life activities, secondary deformities, and pain.[6,7,10] Studies investigating the impact of different motor impairments in children with cerebral palsy (CP) have shown that a loss of SVMC limits motor performance more than other routinely measured impairments such as spasticity and/or contractures.[10-13] Just recently, selective motor control has been listed in the International Classification of Functioning, Disability and Health-Children & Youth (ICF-CY) core sets for children and youth with CP (b7600),[14] underscoring its clinical importance within this patient group. Furthermore, for many ambulatory children with UMN lesions and their caregivers, learning to walk/move 'normally' (i.e. within a physiological pattern or without synergistic mass patterns) is a commonly mentioned rehabilitation goal.[15,16] Although for achieving a 'normalized' walking pattern training of multiple body functions (e.g. balance, strength) is necessary, SVMC plays a major role in performing qualitative good walking movements.[13]

Despite the social and pathophysiological importance of SVMC, evidence about its trainability in children with CP is relatively limited.[17-20] One possible reason for the small number of intervention studies to improve SVMC might be related to the challenges of measuring motor control.[7] In contrast to measuring motor function, which is readily done in numbers by asking 'how fast' or 'how often' a movement is performed, measuring motor control is more difficult as it looks on 'how' the movement is controlled and executed.[1,2,7] Nevertheless, for being able to assess rehabilitation (medical/ therapeutically) induced changes of SVMC the availability of valid, reliable, and responsive SVMC outcome measures are fundamental.

As no systematic review of the psychometric properties of SVMC measures for the lower extremity for children with UMN lesions exists, the purpose of this study was to address this gap. By focusing on the lower extremity, we aimed to extend the observations from a recent systematic review[21] which investigated psychometric properties of tests scoring either compensatory or physiological movements of the upper extremity in children with UMN lesions. Providing an overview of SVMC measures for the lower extremity and the evaluation of the level of evidence of the

psychometric properties of SVMC measures for children with UMN lesion will be useful for clinicians and researchers planning future studies on the trainability of SVMC.

Method

Search Strategy

To identify studies assessing the psychometric properties of outcome measures of SVMC in children with an UMN lesion the following databases were searched without any time limit until July 2016: MEDLINE, EMBASE, CINAHL, PsycINFO, SCOPUS, Cochrane and PEDro. The search strategy included keywords and synonyms for SVMC as well as names of tools previously used to measure SVMC, the population of interest, and a validated search filter for finding studies on measurement properties.[22] Please see Appendix 1 for an example of the search strategy for MEDLINE. In addition, we hand-searched the reference lists of the articles included in the review to identify additional studies.

Study Selection

We used a previously developed proprietary database (Microsoft Access 2010) to systematically enter the data and score the methodological quality of the studies.[23] Inclusion and exclusion criteria were defined in advance. In accordance to the definition of SVMC, stated in the introduction, only papers dealing with selective movement of one joint of the lower extremity or with a primary selective (not synergistic) voluntary multi-joint movement were included. For example, papers dealing with the ankle dorsiflexion during initial contact or investigating pathological synergy patterns during walking (i.e. activation of the m. rectus femoris and m. semitendinosus during swing phase) were included, whereas papers measuring SVMC over the whole gait cycle or during gross motor coordination tasks were excluded. Considering that SVMC comprehends how accurately and smoothly someone can isolate the selection of a particular muscle group, papers describing the measurement of submaximal torque steadiness were included (i.e. ICF body function level b7300 power of isolated muscle activation). However, studies on maximal voluntary contraction were excluded, as patients with impaired SVMC tend to produce maximal force by using mass synergy patterns.[24,25] Furthermore, neuroimaging measures, testing structural and metabolic intactness of the involved underlying neurophysiological structures, or networks involved in SVMC (i.e. ICF body structure levels 1100 CST, primary cortex) were excluded. Only papers dealing with children and youths (3 to 21 years) with UMN lesions were included. This age range was chosen for neurophysiological reasons (e.g. maturation of the corticospinal tract) and practical reasons (e.g. compliance/understanding). Studies with the explicit aim to assess one or more psychometric properties were included, as well as cohort-studies indirectly investigating the psychometric characteristics of an outcome measure, by for instance looking at the difference between neurological intact children and those with UMN lesion. All other forms of indirect evidence (i.e. intervention studies) were excluded. Only manuscripts published in English and German were included for review.

Two reviewers (J.B. and M.vdL.) independently screened all titles and abstracts of the papers. In cases of doubt, the full text article was consulted to decide whether or not the study met the inclusion criteria. A third reviewer was available if no consensus could be achieved.

Quality Evaluation

Evaluation of the methodological quality of the included papers was carried out independently by J.B. and M.vdL by using the 4-point rating scale (‘excellent’, ‘good’, ‘fair’, ‘poor’ or ‘not applicable’) of the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist.[26] The COSMIN-checklist

consists of three domains, namely validity ('The degree to which an (HR-PRO) instrument measures the construct(s) it purports to measure'), reliability ('The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions'), and responsiveness ('The ability of an (HR-PRO) instrument to detect change over time in the construct to be measured').[26] Each domain contains one or more measurement properties. The reviewer selects the measurement properties (COSMIN boxes) evaluated in the study and scores the specific item-lists via the aforementioned ordinal scoring system. The lowest score of all items of the chosen COSMIN box determines the overall methodological quality of the paper. In line with previous COSMIN reviews in the field of neuro-pediatrics, we adopted the overall COSMIN score by omitting the item regarding sample size.[21,23]

To ensure that both raters scored the papers in accordance to the guidelines and to allow other raters to arrive at the same conclusion, the following procedures were established prior to the independent COSMIN rating: raters familiarized with the COSMIN manual and terminology, and discussed the scoring of two papers and established additional rating rules (Appendix 4). Although the COSMIN manual provides general rules for all boxes and items, for some items, the COSMIN rating is still open to subjective interpretation, e.g. 'time interval appropriate'. It is for this reason that COSMIN itself recommends specification of additional rules for individual reviews.[27] If the two reviewers could not agree on a scoring a third reviewer was available.

For the assessment of quality of the measurement properties, the updated criteria suggested by Terwee et al.[28] were applied (Appendix 2). The overall level of evidence for each SVMC measure and each measurement property was evaluated according to the Cochrane Back Review Group Criteria[29] 'strong', 'moderate', 'limited', 'conflicting', 'unknown' (Appendix 3). This overall score was given in relation to the methodological quality of the study and the results of the measurement properties.[26] Again, criteria for sample size were adapted as follows: sample size > 100 subjects of the combined studies was rated as 'strong' [+++ or ---]; sample size between 50-99 was rated as 'moderate' [++ or --]; sample size between 25-59 as 'limited' [+ or -]; and sample size fewer than 25 as 'unknown' [?].[30]

Results

Description of the included studies

The systematic search resulted in 3590 references being identified. Based on the titles and abstracts, 33 papers were included for full-text reading. After applying the inclusion and exclusion criteria, 17 papers were retained for review (Fig. 1).

These 17 papers described the measurement properties of four clinical, ordinal-scaled, assessment tools (Selective Motor Control test (SMC); modified Trost-test (mTrost); Gillette's SMC test; Selective Control Assessment of the lower Extremity (SCALE)) and three laboratory based interval-scaled measurement tools (kinematic measures, electromyography (EMG) and torque steadiness). The majority of studies tested SVMC of the ankle or the knee joint.

The following psychometric properties were evaluated: hypotheses testing/construct validity was assessed in 17 studies, reliability in six (inter-rater n=5; test-retest n=3; intra-rater n=1) and both content and criterion validity in one study. Responsiveness was not evaluated in any study. Most studies tested the SCALE (n=9), followed by studies evaluating torque steadiness measures (n=2), kinematic measures (n=4), and EMG of selected lower limb muscles (n=1). The age of the participants in the studies included for final review ranged from two to 21 years with the exception of one study[31]

where the oldest participant was 28 years. Although this age range was slightly wider than the one set by the inclusion criteria (3-21 years), discussing this issue ended in the common decision for inclusion. As the mean age ranged from 9 years 3 months to 16 years, the youngest and the oldest participant were seen as outliers. Sample size varied from eight to 51 participants. All studies included children with a diagnosis of cerebral palsy. In two cohort studies[32,33], data of children with CP who had undergone a selective dorsal rhizotomy were compared to those of a control group children with CP who had not. Selective dorsal rhizotomy is a neurosurgical procedure, which aims to minimize the limiting influence of spasticity on motor control in children with spastic CP. The comparison of SVMC between children with and without rhizotomy was therefore considered an assessment of the validity of the SVMC tool. Four other cohort studies[31,34-36] investigated the construct validity of the SVMC instrument, by comparing patients with CP versus participants who were neurologically intact. General characteristics and clinical utility for each SVMC measure is summarized in Table I. The methodological quality per measurement property as well as the overall evidence criteria can be seen in Table II.

Hypotheses testing

Of the 17 papers which evaluated construct validity ('hypotheses testing'), 10 papers included clinical assessment tools and seven papers laboratory-based measurement tools (Table I and II).

Nine of the 10 papers regarding clinical assessment tools evaluated construct validity of the SCALE. The modified COSMIN scores of three of these 8 SCALE papers were 'good',[37-39] four were rated as 'fair',[40-42] one as 'excellent',[44] and one as 'poor'.[45] Quality of construct validity was evaluated in accordance to Terwee et al.[29] as 'positive [+]' in eight papers[37-43] and as mixed 'positive/negative [+]/[-]' in one studies[45]. Overall, there was 'moderate positive [++]' evidence[30] for construct validity of the SCALE in terms of: i) its correlation with the GMFCS,[37,44 a] and ii) its proximal-distal concordance (SVMC is more often and/or more severely impaired in distal body parts).[38,44] A 'limited positive [+]' evidence level was given for its validity testing with the Berg Balance Scale[43] and for predicting knee flexion during initial contact during stance phase of gait.[42] Three studies investigated relationships between total limb SCALE scores and knee flexion during swing phase. Two studies[39,40] found significant correlations and one not.[41] Therefore, their level of evidence was rated as 'conflicting [±]'. In relation to the poor quality of the Kusumoto et al. study,[45] who investigated the relationship between the SCALE and knee extensor strength, its level of evidence was rated as 'unknown [?]'. Therefore, this study did not contribute to the overall evidence level of the SCALE."

In the other three SVMC clinical assessment tools construct validity was only evaluated for the Gillette's SMC test. The study of Manikowska and collaeges[31] compared Gillette's knee flexion SMC scores in patients with CP versus participants who were neurologically intact using electromyography. This study received a 'poor' modified COSMIN score, but results were rated as 'positive [+]' in accordance to Terwee et al. [29] The level of evidence[30] was evaluated as 'unknown [?]'.

Seven papers investigated construct validity of SVMC using laboratory-based measurement tools including lower limb kinematics,[32,34,39,40] EMG,[33] and torque steadiness.[35,36] Two kinematic papers[39,40] were already evaluated in relation to the SCALE's construct validity in the previous paragraph. We decided to list them under this sections as well. As there is no gold standard measure for quantifying SVMC, and the papers are cohort studies investigating the correlation between the SCALE and a kinematic measures for SVMC, they could be regarded as studies investigating the construct validity of the SCALE, but also of the kinematic measures, depending on which measure is regarded as more 'established'. As the quality and evidence rating of the two studies[39,40] is the same as presented above, the results will not be repeated

here. Their COSMIN methodological quality was rated as 'fair' and quality of construct validity was rated as 'positive [+]' in three studies[34,35,36] and as mixed 'positive/negative [+]/[-]' in two[32,33]. As the sample size was too small (n<25) in three studies, the evidence level was only scored for one kinematic[34] and one EMG[33] study. As the results of these studies were ambiguous in supporting the construct validity of the SVMC measurement method, a 'conflicting [±]' evidence rating was assigned.
Overall, the methodological quality of the majority of the above mentioned studies was reduced due the absence of *a priori* formulated hypotheses, thereby limiting their COSMIN[27] as well as validity quality[29] scoring.

Content and Criterion Validity
Content and criterion validity were only assessed for the SCALE (Table II).[37,44] COSMIN rating[27] of content validity was considered 'poor'. Although 14 experts were involved on item-agreement for statements about content, administration, and grading of the SCALE, the paper lacked a description whether all items were relevant for the construct or for the population of interest.[37] The quality rating of the results[27] was scored as 'indeterminate [?]', and the evidence as 'unknown [?]'.
The method applied to establish criterion validity of the SCALE was rated as 'excellent'.[44] As the Fugl-Meyer Assessment (item III and IV) measures a similar construct as the SCALE and their correlation exceeded 0.70, the SCALE criterion validity results were rated as 'positive [+]'. Therefore, a 'limited positive [+]' evidence level was given for criterion validity of the SCALE.

Reliability
Reliability was investigated in three (SMC[46,47], mTrost[47], SCALE[37,44]) of the four clinical assessment tools and in two (kinematic[32], torque[35]) of five laboratory-based SVMC tools. The SMC test-retest reliability was tested in two studies.[46,47]. The modified COSMIN[27] rating for SMC inter-rater reliability ranged from 'fair' [47] to 'good'[46]. The methodological quality of inter-rater reliability of the mTrost test was also rated 'fair'. [47] Inter-rater reliability of the SCALE was tested by two studies and scored as 'excellent'[44] and 'good' [37]. The SCALE's intra-rater reliability was further investigated and received a 'good' modified COSMIN score.[44] The methodological quality of test-retest reliability for the kinematic[32] and torque steadiness[35] measure was evaluated as 'good'. Overall, studies assessing inter-rater reliability were rated lowest for COSMIN items describing the statistical procedures (i.e. description of weighted scheme ICC, Kappa). The items regarding the stability of participants between the two or more assessments and the description of test conditions were the most limiting items for the four studies on test-retest and intra-rater-reliability.
Applying the quality criteria[29] for measurement properties revealed 'positive [+]' results for four reliability studies;[32,37,44,46] mixed' positive/negative [+]/[-]' results in three studies,[35,46,47,] and negative [-]' results for inter-rater reliability of the SMC test.
When evaluating the overall evidence level using Cochrane guidelines[30], we found 'moderate positive results [++]' for the inter- and intra-rater reliability of the SCALE.[37,44] 'Moderate negative results [- -]' were evident for the inter-rater reliability of the SMC[46,47] and 'limited negative results [-]' for the m-Trost[47]. Due to the low sample size (n<25) the evidence level of the test-retest reliability of the kinematic[32] and torque steadiness measurement[35] studies was scored as 'unknown [?]'.

Discussion
This review revealed a limited number of psychometric studies investigating SVMC measures in children with UMN lesions. The overall evidence was further limited as 10

out of 17 studies were cohort studies with a limited methodological quality (i.e. 'poor' or 'fair') with the exception of one study[39] which scored 'good' according to modified COSMIN rating guidelines. No study investigated responsiveness, which we consider a crucially important measurement property, especially in context of the ongoing debate about trainability of SVMC. Although we explicitly searched for the whole population of pediatric UMN lesions, only psychometric studies including children with CP were found. Therefore, the results of this review are limited to children with CP and cannot be directly transferred to other children with impaired SVMC (i.e. acquired brain injuries). The chosen age range (2-21 years) for this review, might have been wide when considering developmental issues which are known to influence SVMC (e.g. maturation of CNS function) as well as the importance of the participants' cognitive understanding and motivation for the SVMC measurement procedure/testing. Nevertheless, this age range was chosen due to the overall limited number of studies available for review. Future studies regarding SVMC measures may choose to investigate psychometric properties in separate age groups (i.e. pre CNS maturation 2y-7y and post < 8 years).

The SCALE was the most often investigated assessment tool, in terms of number of studies conducted and in the number of its measurement properties investigated. The SMC and m-Trorst were only rated on reliability, thus lacking evidence on their validity. The Gillette's SMC test was only investigated on its validity, lacking evidence about its reliability. In terms of psychometric quality, the SCALE had the highest level of evidence with a moderate positive level of evidence concerning its inter-rater reliability and its construct validity, and an unknown and limited level of evidence of content and criterion validity, respectively. Another advantage of the SCALE in comparison to the other assessment tools (SMC, m-Trost,) lies in its evaluation of five lower extremity joints rather than one or three joints. In addition, clinical utility (Table I) of the SCALE, as well as of the other SMC assessment tools, was scored high as time, costs and resources are low. However, in terms of limitations, the SCALE's application is limited to children with spastic CP. Furthermore, as its ordinal scoring system relies on the impression of the rater (i.e. therapist, consultant), it is a subjective measurement. Finally, the SCALE's 3 level ordinal scoring system (normal, impaired, and unable) may lack sensitivity to detect certain therapy-induced changes of SVMC. These limitations (spastic CP, ordinal scoring system) also apply for the other three clinical tools.

The construct validity of the kinematic, EMG and torque steadiness was assessed, but none of the papers evaluating these measurement techniques explicitly mentioned that the assessment of validity was an *a priori* objective. Because of this, the formulation of hypotheses was often absent thus diminishing their modified COSMIN score to 'fair'. Only two laboratory based SVMC measures (kinematic, torque steadiness)[32,35] were assessed regarding their test-retest reliability. In terms of psychometric quality as well as clinical utility (see Table I), none of the identified laboratory based measures seem to offer great advantage over the other. The equipment required to record the outcome measures was often customized, making it difficult for other groups (researchers or clinicians) to apply and confirm or extend findings of studies exploring the laboratory based measures using EMG, kinematics, or torque measurements. Furthermore, the measurement procedures appear time consuming and complex in comparison to more routinely applied clinical assessments. Personnel also required extensive training in the application and analysis of these measures (see Table I).

In summary, the results from this systematic review show the limited level of evidence regarding the psychometric properties (reliability and validity) and absence of evidence regarding responsiveness of currently available SVMC measures of the lower extremity in children with UMN lesions.

Methodological Considerations

Low inter-rater agreement when rating the quality of the evidence in systematic reviews (e.g. rating Risk of Bias in Cochrane type reviews) can be an important methodological issue which should be considered when conducting a systematic review.[48] In our review agreement between the raters for all COSMIN items was high. Only five out of 246 items needed further discussion, and none required the rating of a third reviewer. This high agreement was likely the result of the specific rating rules which we established as recommended by the COSMIN group. For example, when scoring the reliability items 4-7 for the SCALE, we decided in advance to score the use of video for the evaluating of the inter- and intra-rater reliability as appropriate, as this allows a discrete evaluation of the scoring system by maintaining the stability of test conditions and patient status as well as saving on time and resources. In contrast, a video approach was not considered to be appropriate for determining test-retest reliability when the stability of the patient is evaluated.

In line with other neuro-pediatric COSMIN reviews [22,24,49] we modified the rating of the sample size item (modified COSMIN score) as the sample size is often limited in clinical neuro-pediatric studies and not comparable with large scale epidemiological healthcare studies using patient-reported outcome measures for which the COSMIN guidelines were initially evaluated. This modified scoring improved the overall rating of all studies with the exception of the construct and content validity score of the studies from Fowler et al.[37] and Zwaan et al.[33] Although we used this modified score, we recommend that future psychometric studies should include a sufficiently large sample size (>30).

Other reasons for scoring poor were the lack of 'a priori' formulated hypotheses (box 'hypothesis testing') and for one study[37] the lack of evaluating each item separately for its content validity (box 'content validity'). While we consider it important that each single question should be evaluated separately for its content validity in a Health Care Questionnaire (where the COSMIN was originally developed for), it could be questioned whether the same rating rules are needed for an assessment tool like the SCALE that consists of a similar procedure repeated for different joints.

Future research directions

The results of this review show, that the SCALE is the most frequently investigated assessment method in the population of CP children and also has the highest quality rating. Its responsiveness to change has not been assessed, but it may be expected that due to its ordinal scoring system its sensitivity to measure changes of SVMC is limited. To improve its sensitivity and simultaneously to benefit from its child-friendly procedure, combining the SCALE with another more sensitive measure appears to be promising. This idea has also been proposed in previous studies.[33,50] While Zwaan et al.[33] found no convincing evidence for detecting extensor and flexor synergies during gait using EMG in children with CP, they reported a significant cross-correlation between extensor synergy activity measured using EMG and the m-Trost test. They concluded that 'EMG measures still may be useful for selective motor control measurement because it measures selectivity at the level of the specific muscles involved, provided the appropriate task is used'.[33] As walking requires selective as well as synergistic movements, we considered it not an appropriate task for the assessment of SVMC. The tasks embedded in the SCALE (isolated single-joint movements) were developed in accordance with the definition of SVMC.[8,37] Combining SCALE's ratings for single-joint movements with EMG recordings would further allow for directly measuring voluntary activation of a muscle even in patients with low muscle strength (manual muscle test grade of 1), where no real joint movement occurs.

Conclusion

This systematic review revealed a limited number of psychometric studies evaluating the validity and reliability of SVMC measures in children with UMN lesions, and no studies evaluating responsiveness. Currently, the SCALE appeared to have the highest level of evidence regarding its reliability and construct validity compared to other clinical and laboratory-based measures of SVMC. However, only by means of reliable, validated, and responsive SVMC tools used in carefully designed intervention studies, it will be possible to provide a scientifically rigorous contribution to the ongoing debate with regard to the possibility to improve SVMC of the lower extremity in children with UMN lesions.

Acknowledgements

This research project was funded by the Swiss National Science Foundation (Grant number 32003B_156646), 'PhysioSwiss' (CH); 'Physiotherapy Science Foundation' (CH) and 'Mäxi-Foundation' (CH). We further acknowledge the Zurich Center for Neuroscience (ZNZ). None of the funders were involved in the study design, data collection, analysis, and manuscript preparation and publication decisions.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest. All ideas and decisions in relation to this study were made independently by the authors.

References

1. Boyce WF, Gowland C, Hardy S, Rosenbaum PL, Lane M, Plews N, Goldsmith C, Russell DJ. Development of a quality-of-movement measure for children with cerebral palsy. Phys Ther 1991; 71(11): 820-828.
2. Rosenbaum PL, Russell DJ, Cadman DT, Gowland C, Jarvis S, Hardy S. Issues in measuring change in motor function in children with cerebral palsy: a special communication. *Phys Ther* 1990; 70(2): 125-31.
3. Campbell SK. Assessment of the child with central nervous system dysfunction. In: Rothstein JM, editor. Measurement in Physical Therapy. New York, NY: Churchill Livingstone Inc; 1985. 207-228.
4. Rothwell JC. Control of human voluntary movement. London: Croom Helm; 1987.
5. Rose J, McGill KC. Neuromuscular activation and motor-unit firing characteristics in cerebral palsy. *Dev Med Child Neurol* 2005; 47: 329-336.
6. Cahill-Rowley K, Rose J. Etiology of impaired selective motor control: emerging evidence and its implications for research and treatment in cerebral palsy. *Dev Med Child Neurol* 2014; 56: 522-528.
7. Fowler EG. Concepts in spasticity and selective motor control in children with spastic cerebral palsy. *Technology Disability* 2010; 22: 207-214.
8. Sanger TD, Chen D, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Definition and classification of negative motor signs in childhood. *Pediatrics* 2006; 118: 2159-2167.
9. Rose J. Selective motor control in spastic cerebral palsy. *Dev Med Child Neurol* 2009; 51(8): 578-579.
10. Vos RC, Becher JG, Voorman JM, Gorter JW, van Eck M, van Meeteren J, Smits D-W, Twisk JW, Dallmeijer AJ. Longitudinal association between gross motor capacity and neuromusculoskeletal function in children and youth with cerebral palsy. *Arch Phys Med Rehabil* 2016; 97(8):1329-133.

11. Østensjø S, Carlberg EB, Vøllestad NK. Motor impairments in young children with cerebral palsy: relationship to gross motor function and everyday activities. *Dev Med Child Neurol* 2004; 9: 580–589.
12. Voorman JM, Dallmeijer J, Knol DL, Lankhorst GJ, Becher JG. Prospective longitudinal study of gross motor function in children with cerebral palsy. Arch *Phys Med Rehabil* 2007; 88: 871-876.
13. Desloovere K, Molenaers G, Feys H, Huenaerts C, Callewaert B, Van de Walle P. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture* 2006; 24: 302-313.
14. Schiariti V, Masse LC. Relevant Areas of Functioning in Children With Cerebral Palsy Based on the International Classification of Functioning, Disability and Health Coding System: A Clinical Perspective. *J Child Neurol* 2014; 30: 216-222.
15. Wu YN,Hwang, M, Ren Y, Gaebler-Spira D, Zhang LQ. Combined passive stretching and active movement rehabilitation of lower-limb impairments in children with cerebral palsy using a portable robot. *Neurorehabil Neural Repair* 2011; 25(4): 378–85.
16. Beveridge B, Feltracco D, Struyf J, Strauss E, Dang S, Phelan S, Wright FV, Gibson BE. "You gotta try it all": Parents' Experiences with Robotic Gait Training for their Children with Cerebral Palsy. *Phys Occup Ther Pediatr* 2015; 35: 327-341.
17. Rios DC, Gilbertson T, McCoy SW. NeuroGame Therapy to improve wrist control in children with cerebral palsy: a case series. *Dev Neurorehabil* 2013; 16: 398-409.
18. Sukal-Moulton T, Clancy T, Zhang LQ, Gaebler-Spira D. Clinical application of a robotic ankle training program for cerebral palsy compared to the research laboratory application: does it translate to practice? *Arch. Phys. Med. Rehabil.* 2014; 95: 1433-1440.
19. Burdea GC, Cioi D, Kale A, Janes WE, Ross SA, Engsberg JR. Robotics and gaming to improve ankle strength, motor control, and function in children with cerebral palsy-a case study series. *IEEE Trans Neural Syst Rehabil Eng 2013;* 21: 165-173.
20. Van Doornik J, Kukke S, Mcgill K, Rose J, Sherman-Levine S, Sanger TD. Oral Baclofen Increases Maximal Voluntary Neuromuscular Activation of Ankle Plantar Flexors in Children With Spasticity Due to Cerebral Palsy. *J Child Neurol* 2008: 635-639.
21. Gerber CN, Labruyere R, van Hedel HJ. Reliability and Responsiveness of Upper Limb Motor Assessments for Children With Central Neuromotor Disorders: A Systematic Review. *Neurorehabil Neural Repair. 2016; 30:19-39*.
22. Terwee CB, Jansma EP, Riphagen II, De Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual. Life Res* 2009; 18: 1115-1123.
23. Ammann-Reiffer C, Bastiaenen CHG, de Bie R, van Hedel HJ. Measurement Properties of Gait-Related Outcomes in Youth With Neuromuscular Diagnoses: A Systematic Review. *Phys Ther* 2014; 94: 1067-1082.
24. Damiano DL, Dodd K, Taylor NF. Should we be testing and training muscle strength in cerebral palsy ? *Dev Med Child Neurol* 2002; 44: 68-

25. Tedroff K, Knutson LM, Soderberg GL. Synergistic muscle activation during maximum voluntary contractions in children with and without spastic cerebral palsy. *Dev Med Child Neurol* 2006; 48: 789-796.
26. Mokkink LB, Terwee CB. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments : an international Delphi study. *Qual Life Res* 2010; 19: 539-549.
27. Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL, Knol DL, Bouter LM, de Vet HC. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol* 2010;22:10:82.
28. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.
29. Van Tulder M, Furlan A, Bombardier C, Bouter L; Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003; 28: 1290-1299.
30. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthritis Cartilage* 2012; 20:1548-62.
31. Manikowska F, Chen B, Jozwiak M, Lebiedowska MK.Assessment of selective motor control in clinical Gillette's test using electromyography. *Europ J Phys Rehabil Med* 2016; 52(2): 176-185.
32. Engsberg JR, Ross SA, Park TS. Quantifying Active Ankle Range of Motion in Cerebral Palsy Following Selective Dorsal Rhizotomy. *J Appl Biomech* 2004; 20: 103-111.
33. Zwaan E, Becher JG, Harlaar J. Synergy of EMG patterns in gait as an objective measure of muscle selectivity in children with spastic cerebral palsy. *Gait Posture* 2012; 35: 111-115.
34. Engsberg JR, Van Emmerik RE, Ross SA, Collins DR. Use of relative phase as a measure of motor control at the ankle in persons with cerebral palsy: a preliminary study. *J Appl Biomech* 2008; 24: 382-390.
35. Bandholm T, Rose MH, Sløk R, Sonne-Holm S, Jensen BR. Ankle torque steadiness is related to muscle activation variability and coactivation in children with cerebral palsy. *Muscle Nerve* 2009; 40: 402-410.
36. Arpin DJ, Stuberg W, Stergiou N, Kurz MJ. Motor control of the lower extremity musculature in children with cerebral palsy. *Res Dev Disabil* 2013; 34: 1134-1143.
37. Fowler EG, Staudt L, Greenberg MB, Oppenheim WL. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. *Dev Med Child Neurol* 2009; 51(8): 607-614.

38. Fowler EG, Staudt L, Greenberg MB. Lower-extremity selective voluntary motor control in patients with spastic cerebral palsy: Increased distal motor impairment. *Dev Med Child Neurol* 2010; 52: 264-269.
39. Goldberg EJ, Requejo PS, Fowler EG. Joint moment contributions to swing knee extension acceleration during gait in individuals with spastic diplegic cerebral palsy. *Gait Posture* 2011; 33: 66-70.
40. Fowler EG, Goldberg EJ. The effect of lower extremity selective voluntary motor control on interjoint coordination during gait in children with spastic diplegic cerebral palsy. *Gait Posture* 2009; 29: 102-107.
41. Rha DW, Cahill-Rowley K, Young J, Torburn L, Stephenson K, Rose J. Biomechanical and clinical correlates of swing-phase knee flexion in individuals with spastic Cerebral Palsy who walk with flexed-knee gait. *Arch Phys Med Rehabil* 2015; 96: 511-517.
42. Rha DW, Cahill-Rowley K, Young J, Torburn L, Stephenson K, Rose J. Biomechanical and clinical correlates of stance-phase knee flexion in persons with spastic Cerebral Palsy. *PM R.* 2016; 8:11-18.
43. Lim H. Correlation between the selective control assessment of lower extremity and pediatric balance scale scores in children with spastic cerebral palsy. *J Phys Ther Sci* [2015; 27: 3645–3649.
44. Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy. *Dev Med Child Neurol* 2015; 12805.
45. Kusumoto Y, Takaki K, Matsuda T, Nitta O. Relation of selective voluntary motor control of the lower extremity and extensor strength of the knee joint in children with spastic diplegia. *J Phys Ther Sci* 2016; 28: 1868–1871.
46. Löwing K, Carlberg EB. Reliability of the selective motor control scale in children with cerebral palsy. *Adv Physiother* 2009; 11: 58-63.
47. Smits D-W, van Groenestijn AC, Ketelaar M, Scholtes V, Becher JG, Gorter JW. Selective motor control of the lower extremities in children with cerebral palsy: inter-rater reliability of two tests. *Dev Neurorehabil* 2010; 13: 258-265.
48. Jørgensen L, Paludan-Müller AS, Laursen DR, Savović J, Boutron I, Sterne JA, Higgins JP, Hróbjartsson A. Evaluation of the Cochrane tool for assessing risk of bias in randomized clinical trials: overview of published comments and analysis of user practice in Cochrane and non-Cochrane reviews. *Syst Rev.* 2016; 10- 5: 80.
49. Lennon N, Thorpe D, Balemans AC. The clinimetric properties of aerobic and anaerobic fitness measures in adults with cerebral palsy : A systematic review of the literature. *Res Dev Disabil* 2015; 45-46: 316-32.
50. Dobson F. Assessing selective motor control in children with cerebral palsy. *Dev Med Child Neurol* 2010; 52: 409-410.

Table I: General characteristic, psychometric properties, and clinical utility of SVMC measures

General Characteristic				Psychometric Properties		Clinical Utility				
Instrument	**Body part tested**	**Scale**	**Outcome**	**Reliability (BoxB)***	**Validity (BoxD,F,H)***	**Clarity of instructions**	**Format**	**Qualifi-cations**	**Time**	**Costs**
b7600 (control of simple voluntary movements)										
Selective Motor Control	ankle	ordinal (0-4)	max 4 points leg score	**Inter-rater: *fair-good*** Lowing and Carlberg (2009); Smits et al. (2010)[46,47] **Test-Retest: *fair*** Lowing and Carlberg (2009)[46]	No study	excellent	assessment	not addressed	5 min	none
modified Trost test	ankle, knee, hip	ordinal (0-2)	max 8 points leg score	**Inter-rater: *fair*** Smits et al. (2010)[46]	No study	excellent	assessment	recommended	15 min	none
Gillett's SMC test	for all joints	ordinal (0-2)	max 2 points per joint	No study	**Construct** (only for knee-flexion): Manikowska et al. (2016)[31]	excellent	assessment	recommended	15 min	none
Selective Control Assess-ment of the Lower Extremity	toes, STJ, ankle, knee, hip	ordinal (0-2)	max 10 points leg score	**Inter-rater: *good-excellent*** Fowler et al. (2009)[37]; Balzer et al. (2014)[44] **Intra-rater: *excellent*** Balzer et al. (2014)[44]	**Content: *poor*** (Fowler et al. 2009)[37] **Construct: fair-*good*** Fowler et al. (2009)[37]; Balzer et al. (2015)[44]; Kusumoto et al. (2016)[45]; Lim et al. (2015)[43] **Cohort: fair-*good*** Fowler et al. (2010)[38]; (*)Fowler and Goldberg (2009)[40]; (*)Goldberg et al. (2011)[39], Rha et al. (2015 and 2016)[41,42] **Criterion: *excellent*** Balzer et al. (2015)[44]	excellent	assessment	recommended	15 min	none
Kinematic	ankle	interval	end range ROM DE and PF	**Test-Retest: *good*** Engsberg et al. (2004)[32]	**Cohort: *fair*** Engsberg et al. (2004)[32]	excellent	video recording system	required	not stated	high
			SD relative phase measure (°)	No study	**Cohort: *fair*** Engsberg et al. (2008)[34]	excellent	video recording system	required	not stated	high
	swing phase knee	interval	hip-knee-angle-diagram (minimum relative phase)	No study	(*)Fowler and Goldberg (2009)[40]	adequate	3D gait analysis kinematic	required	not stated	high
		interval	knee extension acceleration	No study	(*)Goldberg et al. (2011)[39]	adequate	3D gait analysis kinematic	required	not stated	high
EMG	ankle, knee, hip	interval	EMG cross correlation: thigh and extensor synergies	No study	**Cohort: *fair*** Zwaan et al. (2012)[33]	adequate	EMG (plus three-dimensional gait analysis)	required	not stated	high

ROM: range of motion, *DE*: Dorsal Flexion; *PF:* Plantar Flexion; *SD:* Standard Deviation; (*) studies are listed both for the SCALE and kinematic measures; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] Table is separated for ICF functions (b7600 and b7300).

Table I. (continued)

General Characteristic				Psychometric Properties		Clinical Utility				
Instrument	**Body part tested**	**Scale**	**Outcome**	**Reliability (BoxB)***	**Validity (BoxD,F,H)***	**Clarity of instructions**	**Format**	**Qualifi-cations**	**Time**	**Costs**
b7300 (power of isolated muscles and muscle groups, muscle activation)										
Torque steadiness & EMG	ankle	interval	SD of target torques: DE: 0.1 and 0.3 Nm/kg body weight PF: 0.3 and 0.5 Nm/kg body weight	**Test-Retest: *good*** Bandholm et al. (2009)[35]	**Cohort: f*air*** Bandholm et al. (2009)[35]	excellent	customized equipment, isometric torque measure (plus EMG)	required	not stated	high
	hip, knee, ankle	interval	SD of [torque/mean torque] *100)	No study	**Cohort: f*air*** Arpin et al. (2013)[36]	excellent	isokinetic dynamometer (plus EMG)	required	not stated	high

ROM: range of motion, *DE*: Dorsal Flexion; *PF:* Plantar Flexion; *SD:* Standard Deviation; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] Table is separated for ICF functions (b7600 and b7300).

Table II: Summary study details, COSMIN-, quality and evidence-rating

			Study Population			Construct validity: Hypothesis Testing (BoxF) (I)					
ICF	**Instrument**	**Study**	**Diagnose**	**Age (SD)**	**n**	**Design**	**Results**	**COSMIN***	**Modified COSMIN****	**Quality Criteria**	**Overall evidence**
b7600 (control of simple voluntary movements)	**Selective Control Assessment of the Lower Extremity (SCALE)**	Fowler et al. (2009)[37]	spastic CP GMFCS I: 10; II:12; III: 19; IV:10	mean 11y,11m (4y,9m)	51	Construct validity: GMFCS	SCALE and GMFCS sig. rho: -0.83, sig	good	good	+	+ + (SCALE – GMFCS)
		Fowler et al. (2010)[38]	spastic CP GMFCSI: 10; II: 10; III:18; IV:9	mean age 11y9m (4y8m)	48	Cohort study: pathophysiology (increased distal impairment, in relation to impaired corticospinal tracts)	- sig. differences in SCALE joint scores overall joint - sig. differences between all pairs of hip, knee; ankle - *proximal to distal concordance*	fair	good	+	+ + (SCALE – joint)
		(*)Fowler and Goldberg (2009)[40]	spastic CP GMFCS I-IV	range 6y-21y	15	Cohort study: SCALE and correlation interjoint-coordination hip-knee angle diagrams during swing phase of gait	- higher SCALE scores normal hip-knee-angle-diagram; ROM increased - correlation total SCALE score and MRP: sig. rho 0.83 - correlation total SCALE and velocity: rho 0.66	poor	fair	+	± (SCALE – knee swing phase)
		(*)Goldberg et al. (2011)[39]	spastic CP GMFCS I:5; II:8; III:5	mean age 13y8m (7y2m)	18	Cohort study: correlation SCALE limb score and total swing joint contribution/swing extension and acceleration	- sig. PCC: 0.85 - as SCALE scores increased, the swing joint movements provided less resistance to knee extension - SCALE 0-4 = simultaneous hip and knee flexion = diminished knee extension acceleration	poor	good	+	
		Rha et al. (2015)[41]	spastic CP GMFCS I-III	mean age 10y1m	34	Cohort study: regression: SCALE; contracture knee and ankle, gait-kinematics and muscle-tendon length	- SCALE scores did not sig. correlated with the magnitude nor the timing of peak knee flexion during swing phase	fair	fair	-	
		Rha et al. (2016) [42]				Cohort study: regression: SCALE; minimum knee flexion angle at initial contact, gait-kinematics and muscle-tendon length	- lower SCALE scores sig. correlated (rho -0.53) with knee flexion during initial contact - SCALE scores (p=.001) and delayed timing in peak knee flexion during swing (p=.026) independently predicted knee flexion during initial contact	fair	fair	+	+
		Balzer et al. (2015)[44]	spastic CP GMFCS I: 23; II:5; III:8; IV: 3	mean age 12y6m (3y7m)	39	Construct validity: GMFCS, limb distribution correlations: MAS MMT	- sig. differences between GMFCS levels (I vs II), most joint pairs, less and more affected leg - sig. correlation: MMT (rho: 0.88) MAS (rho: -0.55)	fair	excellent	+	+ + (SCALE – GMFCS) (SCALE-joint)
		Lim et al. (2015)[43]	spastic CP GMFCS I: 5; II:8; III:7; IV: 3	mean age 9y3m (2y3m)	23	Construct validity: SCALE is a valid tool to predict the PBS	sig. differences between hemi- and diplegia, - sig. correlation for total and itmes SCALE vs PBS (rho: 0.62-0.79)	poor	fair	+	+
		Kusumoto et al. (2016)[45]	spastic CP GMFCS I: 12; II:19; III:9	mean 13y3m (3y4m)	40	Construct validity: correlation between SCALE and knee extensor strength	- sig. differences in SCALE scores more and less affected leg, no differences for knee extensor strength - inverse sig. correlation SCAL E and knee extensor strength (rho: 0.4)	poor	poor	+/-	?

CP: cerebral palsy*; GMFCS:* Gross Motor Function Classification Level*; rho:* spearman's rank correlation coefficient; *ROM:* Range Of Motion; *MRP:* Minimum Relative Phase (measurement of interjoint coordination between the hip and the knee)*; PCC:* Pearson product-moment correlation coefficient; *MAS:* Modified Ashworth Scale; *MMT:* Manual Muscle Testing; *PBS:* Pediatric Berg Balance Scale; (*) studies are listed both for the SCALE and kinematic measures; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] ** modified COSMIN: omitting the sample size item boxB03; *Quality Criteria:* according to Terwee et al. [28] (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; *Overall evidence;* according to Cochrane van Tulder et al. 2003[29] (Appendix 3): strong (+++/---), moderate (++/--), limiting (+/-) = positive/ negative results; ± conflicting results; ? = unknown results.

Table II. (continued)

			Study Population			Construct validity: Hypothesis Testing (BoxF) (II)					
ICF	**Instrument**	**Study**	**Diagnose**	**Age (SD)**	**n**	**Design**	**Results**	**COSMIN***	**Modified COSMIN****	**Quality Criteria**	**Overall evidence**
b7600 (control of simple voluntary movements)	**Gillett's SMC test**	Mani-kowska et al. (2016)[31]	CP GMFCS I-III	CP mean age 15y (5y5m); TD 22y (1y5m)	CP 23 TD 19	Observational study: CP muscle activity (EMG) during selective knee flexion versus neuro-logical intact controls	sig. differences between CP and Control, and between CP and different levels of SVMC (0 vs 2)	poor	poor	+	?
	Kinematic	Engsberg et al. (2008)[34]	spastic CP bilateral GMFCS I: 5; II: 8; III:1	CP mean age 16y (10y); TD age 15y (9y)	CP 29 TD 15	Cohort study: relative phase = quantification of the relative timing between a pair of oscillators at the same frequency	- sig. differences between CP (high SD) and TD (low SD) children in DE / PF movements - no sig. differences for antiphase movement	poor	fair	+	±
		Engsberg et al. (2004)[32]	CP pre post SDR spastic CP	CP mean age 7y; TD mean age 7y2m	CP pre post SDR 12 CP 14 TD 20	Cohort study: differences between active ROM ankle in CP children pre and post SDR and vs TD	- sig. increase in active full ROM and active DE (but not PF) pre vs post SDR - no sig. differences in CP only group - sig. smaller ROM CP vs TD	poor	fair	+/-	(sample size)
		(*)Fowler and Goldberg (2009)[40]	spastic CP GMFCS I-IV	range 6y-21y	15	Cohort study: correlation interjoint-coordination hip-knee angle diagrams during swing phase of gait with total limb SCALE score	- higher SCALE scores normal hip-knee-angle-diagram; ROM increased - correlation MRP and SCALE score: sig. rho 0.83	poor	fair	+	+
		(*)Goldberg et al. (2011)[39]	spastic CP GMFCS I:5; II:8; III:5	mean age 13y8m (7y2m)	18	Cohort study: correlation knee extension acceleration during swing and total limb SCALE score	- sig. PCC: 0.85 - as SCALE scores increased, the swing joint movements provided less resistance to knee extension - SCALE 0-4 = simultaneous hip and knee flexion = diminished knee extension acceleration	poor	good	+	+
	EMG (thigh- and extensor-synergies)	Zwaan et al. (2012)[33]	CP post SDR , (CP; GMFCS I-III)	CP mean age 6y5m; TD range 6-11y	CP post SDR 39 CP 38 TD 30	Cohort study: CP vs TD, correlation between synergies pattern (EMG), gait profile, mTrost, GMFM	- extensor synergy: sig. differences CP (0.95) vs TD (0.77) - thigh synergy: only sig. differences t for comfortable speed CP (0.94) vs TD (0.95) - no strong correlation between mTrost and gait EMG /GMFM	fair	fair	+/-	±

CP: cerebral palsy*; TD:* Typically Developed children; *GMFCS:* Gross Motor Function Classification Level*; SD:* Standard Deviation; *DE*: Dorsal Extension; *PF*: Plantar Flexion; *SDR:* Selective Dorsal Rhizotomy; *ROM:* Range Of Motion; *GMFM:* Gross Motor Function Measure; *CV:* average coefficient of variation; *rho:* spearman's rank correlation coefficient; (*) studies are listed both for the SCALE and kinematic measures; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] ** modified COSMIN: omitting the sample size item boxB03; *Quality Criteria:* according to Terwee et al. [28] (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; *Overall evidence;* according to Cochrane van Tulder et al. 2003[29] (Appendix 3): strong (+++/---), moderate (++/--), limiting (+/-) = positive/ negative results; ± conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).

Table II. (continued)

			Study Population			Construct validity: Hypothesis Testing (BoxF) (II)					
ICF	**Instrument**	**Study**	**Diagnose**	**Age (SD)**	**n**	**Design**	**Results**	**COSMIN***	**Modified COSMIN****	**Quality**	**ICF**
b7300 (power of isolated muscles and muscle groups, muscle activation)	**Torque steadiness**	Bandholm et al. (2009)[35]	CP unilateral, GMFCS I: 13, II:1	CP mean age 11y (1y); TD mean age 11y (1y)	CP 14 TD 14	Cohort study: comparison CP and TD torque steadiness DE/PF	- sig. reduction of torque steadiness in CP (SD and CV were higher) vs TD - DE was most affected - sig. greater antagonist-agonist activation ratio and muscle activation variability in CP vs TD - DE torque steadiness correlated with level of PF coactivation (r = 0.597) and DE antagonist-agonist activation ratio (r= 0.832)	poor	fair	+	(sample size)
		Arpin et al. (2013)[36]	CP bilateral and unilateral, GMFCS I: 7; II:5; III:3	CP age mean 14y2m (7m); TD age mean 14y1m (7m);	CP 15 TD 15	Cohort study: comparison CP and TD torque steadiness DE/PF	- sig. greater CV at the ankle in CP vs TD - CP: sig. greater variability at ankle, then knee and hip - CP more regular steady torque patterns vs TD	poor	fair		(sample size)
			Study Population			**Content Validity (BoxD) / Criterion Validity (BoxH)**					
b7600 (control of simple voluntary movements)	**Selective Control Assessment of the Lower Extremity (SCALE)**	Fowler et al. (2009)[37]	spastic CP, GMFCS I: 10; II:12; III: 19; IV:10	mean 11y,11m (4y,9m)	51	Content validity: 14 experiences clinicians	mean agreement: 91.9% (range 71.4–100%) for content, administration, and grading	poor/	poor	?	?
		Balzer et al. (2015)[44]	spastic CP GMFCS I: 23; II:5; III:8; IV: 3	mean age 12y6m (3y7m)	39	Criterion validity: correlation SCALE and FMA item III & IV	sig. correlation: FMA (rho:0.88)	fair	excellent	+	++

CP: cerebral palsy*; TD:* Typically Developed children; *GMFCS:* Gross Motor Function Classification Level*; SD:* Standard Deviation; *DE*: Dorsal Extension; *PF*: Plantar Flexion; *SDR:* Selective Dorsal Rhizotomy; *ROM:* Range Of Motion; *GMFM:* Gross Motor Function Measure; *CV:* average coefficient of variation; *rho:* spearman's rank correlation coefficient; *FMA:* Fugl-Meyer Assessment; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] ** modified COSMIN: omitting the sample size item boxB03; *Quality Criteria:* according to Terwee et al. [28] (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; *Overall evidence;* according to Cochrane van Tulder et al. 2003[29] (Appendix 3): strong (+++/---), moderate (++/--), limiting (+/-) = positive/ negative results; ± conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).

Table II. (continued)

			Study Population			Reliability (Box B)					
ICF	**Instrument**	**Study**	**Diagnose**	**Age (SD)**	**n**	**Design**	**Results**	**COSMIN***	**Modified COSMIN****	**Quality Criteria**	**Overall evidence**
b7600 (control of simple voluntary movements)	**Selective Motor Control test (SMC)**	Lowing and Carl-berg (2009)[46]	CP bilateral n=30: unilateral n=10; GMFCS I:13; II:12; III:10; IV:3 and V:2	median age 7y; range 3y-16y	29	Test-Retest (1-60 days)	Kw ankle DE = 0.88-1; RV: 0 - 0.005; RP: -0.11 - 0; RC: 0 - 0.05 (right & left leg scores)	poor	fair	+	+ (SMC Test-Retest)
					40	Interrater (3 PTs)	Kw ankle DE: 0.58-0.77 RV: 0.000-0.030; RP: -0.09 - 0.11; RC: -0.26 - 1.16 (right & left leg scores)	fair	good	+/-	- - (SMC interrater)
	Selective Motor Control test (SMC)	Smits et al. (2010)[47]	spastic CP	mean age 6y,5m (12m)	21	Interrater (1 PT and 1 Dr) 1h in-between SMC and mTrost	K ankle DE: 0.55 95% IC: 0.36-0.74 (leg total score)	poor	fair	-	
	Modified Trost (m-Trost)						K ankle DE: 0.65; 95% IC: 0.47-0.84 K knee EXT: 0.69; 95% IC: 0.49-0.88 K hip ABD: 0.57; 95% IC: 0.37-0.78 K hip FLEX: 0.71; 95% IC: 0.51-0.91	poor	fair	+/-	- (m-Trost interrater)
	Selective Control Assessment of the Lower Extremity (SCALE)	Fowler et al. (2009)[37]	spastic CP	mean age 12y,3m (5y5m)	20	Interrater (2 groups: 3 PT and 3 Dr)	ICC: 0.88-0.91; 95% CI: 0.69-0-97 (right & left leg scores)	poor	good	+	+ + (SCALE interrater)
		Balzer et al. (2015)[43]	spastic CP GMFCS I: 23; II:5; III:8; IV: 3	mean age 12y6m (3y7m)	38	Interrater (2 PTs via Video)	ICC:0.91-0.94 (less & more affected leg); MMD: 1.88-1.92	fair	excellent	+	
						Intrarater (1 rater via video, 6 weeks)	ICC:0.95-0.96 (less & more affected leg); MMD: 1.79-1.96	fair	good	+	+ + (SCALE intrarater)
	Kinematic	Engsberg et al. (2004)[32]	Spastic CP	not stated (pilot-study)	8	Test-Retest (8 weeks)	sig. PCC: ankle PF: 0.77; ankle DE: 0.94; total sagittal range: 0.93 mean differences of: max PF: 1.6°; DE = 1.1°; total range = 2.4°	poor	good	+	? (sample size) (Test-Retest)
b7300 (power of isolated muscles and muscle groups, muscle activation)	**Torque steadiness**	Bandholm et al. (2009)[35]	CP	not stated (pilot-study)	7	Test-Retest (1 day)	ICC ankle PF:0.72; CV < 19% ICC ankle DE: 0.31; CV 25%	poor	good	+/-	? (sample size) (Test-Retest)

CP: cerebral palsy*; GMFCS:* Gross Motor Function Classification Level*; Kw:* Kappa weighted; *DE:* Dorsal Flexion*; PF:* Plantar Flexion; *RV:* Relative rank Variance (random disagreement); *RP:* Relative Position; *RC:* Relative Concentration (systematic disagreement); *PT*: Physiotherapist*; Dr:* Physician; *K:* Kappa; *95% IC*: 95% Confidence interval; *EXT:* Extension; *ABD:* Abduction; *FLEX:* Flexion; *ICC:* Infraclass Correlation Coefficient; *MMD:* Minimal detectable change; *PCC:* Pearson product-moment correlation coefficient; *CV:* average coefficient of variation, *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; [26] ** modified COSMIN: omitting the sample size item boxB03; *Quality Criteria:* according to Terwee et al. [28] (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; *Overall evidence;* according to Cochrane van Tulder et al. 2003[29] (Appendix 3): strong (+++/---), moderate (++/--), limiting (+/-) = positive/ negative results; ± conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).